Supplementary Information

Generalized Discrete-Time Markov Models of Infectious Disease Spread

S1 Calculating the Transition Probability Matrix for the Markov chain $\{(\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t)) : t = 1, 2, \ldots\}$

Let P denote the transition probability matrix for the Markov chain $\{(\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t)) : t = 1, 2, \ldots\}$, and $\mathrm{P}(i, j)$ denote the (i, j) entity. To calculate P, suppose first that the states $(\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t))$ are sorted in increasing order of first $\Theta_{C_1}(t)$, second $\Theta_{C_2}(t)$, and so on; that is, if in matrix P the state $(\Theta_{C_1}(t), \Theta_{C_2}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{i_1}^{C_1}, \theta_{i_2}^{C_2}, \ldots, \theta_{i_M}^{C_M})$ is located in row (column) i, and the state $(\Theta_{C_1}(t), \Theta_{C_2}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{j_1}^{C_1}, \theta_{j_2}^{C_2}, \ldots, \theta_{j_M}^{C_M})$ is located in row (column) j, then $i \geq j$ if and only if $\theta_{i_1}^{C_1} \geq \theta_{j_1}^{C_1}, \theta_{i_2}^{C_1} \geq \theta_{j_2}^{C_1}, \ldots, \theta_{i_M}^{C_M} \geq \theta_{j_M}^{C_M}$.

Now, let's consider a model in which the transition probability from one state to another should be set to zero in the matrix P. For instance, when modeling a SIR-type disease spreading in a closed population (where those that are infected cannot return to the susceptible class), the transition matrix P will be lower-triangular, since $\Theta_S(t + \Delta t) \leq \Theta_S(t)$ (i.e. the proportion susceptible in the next state cannot be greater than the proportion susceptible in the current state); hence, $\mathrm{P}(i, j) = 0$ if $i < j$.

Before proceeding to the calculations, it should be pointed out that if in a Markov chain $\{(\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t)) : t = 1, 2, \ldots\}$, the transition between any two states is possible (even with infinitesimal probability), the approximation error described in §3 does not occur (for instance, in the case of an SIS model in §4.1. Also, any form of inequality restriction on the state transitions (e.g., $\Theta_S(t + \Delta t) \leq \Theta_S(t)$) leads to a transition matrix consisting of several lower-triangular sub-matrixes (as for the SIR model for which the transition matrix is a complete lower-triangular matrix). Therefore, here we only describe the calculation of lower-triangular transition matrixes.

If the transition probability (33) turns out to be greater than zero for the states corresponding to entity $(i, j), i < j$, of the lower-triangular transition matrix P, then an intuitive approximation

Table 1. Calculating lower-triangular probability matrix P

If transition from state $(\Theta_{C_1}(t), \Theta_{C_2}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{i_1}^{C_1}, \theta_{i_2}^{C_2}, \ldots, \theta_{i_M}^{C_M})$ in row i to state $(\Theta_{C_1}(t), \Theta_{C_2}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{j_1}^{C_1}, \theta_{j_2}^{C_2}, \ldots, \theta_{j_M}^{C_M})$ in column j cannot occur:
1. Form the support $\Omega_{X(t)}$ given by (9), for $X_{C_1}(t) = \left\lfloor N\theta_{i_1}^{C_1} \right\rfloor, \ldots, X_{C_M}(t) = \left\lfloor N\theta_{i_M}^{C_M} \right\rfloor$.
2. Form
$$\Omega_{\Theta(t)} = \{(x_1, \ldots, x_M) \in \mathbb{N}^M \mid \left\lceil Nb_{j_1-1}^{C_1} \right\rceil \leq x_1 \leq \left\lfloor Nb_{j_1}^{C_1} \right\rfloor, \ldots, \left\lceil Nb_{j_M-1}^{C_M} \right\rceil \leq x_M \leq \left\lfloor Nb_{j_M}^{C_M} \right\rfloor\},$$
where the control points $\{b_{j_1-1}^{C_1}, b_{j_1}^{C_1}\}, \ldots, \{b_{j_M-1}^{C_M}, b_{j_M}^{C_M}\}$ correspond to the state $(\Theta_{C_1}(t), \Theta_{C_2}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{j_1}^{C_1}, \theta_{j_2}^{C_2}, \ldots, \theta_{j_M}^{C_M})$.
3. If $\Omega_{X(t)} \cap \Omega_{\Theta(t)} \neq \emptyset$ then
$$\mathrm{P}(i,i) = \mathrm{P}(i,i) + \Pr\{(\Theta_{C_1}(t+\Delta t), \ldots, \Theta_{C_M}(t+\Delta t)) = (\theta_{j_1}^{C_1}, \ldots, \theta_{j_M}^{C_M}) \mid (\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{i_1}^{C_1}, \ldots, \theta_{i_M}^{C_M})\},$$
where the second term in the right-hand side is given by Eq.33.
Else
$$\mathrm{P}(i,j) = \Pr\{(\Theta_{C_1}(t+\Delta t), \ldots, \Theta_{C_M}(t+\Delta t)) = (\theta_{j_1}^{C_1}, \ldots, \theta_{j_M}^{C_M}) \mid (\Theta_{C_1}(t), \ldots, \Theta_{C_M}(t)) = (\theta_{i_1}^{C_1}, \ldots, \theta_{i_M}^{C_M})\},$$
given by Eq.33.
End If

is to add this probability to the diagonal entity (i, i) (remember that $\mathrm{P}(i, j)$ should set to zero for $i < j$). Table 1 summarizes the steps in calculating a transition matrix that is lower triangular.

To improve the efficiency of the procedure described in Table 1, when the number of trials is large enough a Normal distribution can be used to approximate the Binomial distributions. By a commonly used rule, a normal approximation is appropriate only if everything within κ standard deviations of its mean is within the range of possible values; that is, to approximate the Binomial (n, p), Normal $(np, np(1-p))$ can be used if:

$$np \pm \kappa\sqrt{np(1-p)} \in [0, n].$$

Setting $\kappa = 3$ often results in an accurate approximation; however, in order to increase the accuracy of the SIR model presented in §6, we used $\kappa = 4$ to determine whether the Normal approximation was appropriate.